Book Reviews

Introduction to Bioinformatics. By Teresa K. Attwood and David J. Parry-Smith. Addison Wesley Longman Limited. 1999. ISBN 0-582-32788-1. 240 pages. Price £19.99.

This is intended as a senior undergraduate/M.Sc. textbook. It is a slim, readable paperback with attractive and clear diagrams. It concentrates mainly on protein sequence analysis, and 'the essence of sequence analysis is the detection of homologous sequences by means of routine database searches'. Sequence analysis tools are mostly described generically (the book is not a substitute for any program or package manual), but there is a strong emphasis on access *via* the World Wide Web. The final chapter presents an overall strategy for characterizing an unknown sequence, and this Chapter is accompanied by an interactive WWW tutorial at
http://www.bioinf.man.ac.uk/dbbrowser/bioactivity/prefacefrm.html
The tools at this website (for translation, database searching, etc.) work reasonably well. However, the site makes extensive use of 'frames', with the irritating and by now well-known consequence that users can frequently get into dead ends from which the 'back' button and the 'history' menu cannot rescue them. The tools at this site are too simplified for use on real problems, and many course organisers will prefer to use one of the general purpose sites (such as those at the European Bioinformatics Institute or Baylor College of Medicine) as the basis for tutorial exercises.

The authors sketch very lightly the technical details of the algorithms used for database searching, and the underlying theory. However, I am sure that this is the right approach in a textbook at this level, and they correctly emphasize throughout the limited value that computer methods have except as an adjunct to experimental studies. The best parts of the book are those which show how a biologist's insight into the systems he works with can be extended by identifying sequence similarities in other organisms or with other, related, functions.

It is obviously important that a book which is short on underlying theory and statistics should be particularly careful in offering general advice and rules of thumb. There are some rather surprising (considering the authors' combined experience) oversights in this respect. The current accuracy of protein secondary-structure prediction is very seriously underestimated, and no useful guidance to programs is provided. It has been generally appreciated for many years that the significance of protein sequence similarity cannot be inferred from the overall percentage identity – the length over which this extends is equally important. The only protein similarity matrix presented (Dayhoff 250 PAM) should never be the first choice for sequence database searching. Overall, the book can be recommended for its target audience, though the speed of change in this field is still so fast that a book can never produce as up to date an introduction as a short, 'hands-on' course.

ANDREW COULSON
Biocomputing Research Unit,
Institute of Cell and Molecular Biology,
University of Edinburgh

Crystallisation of Biological Macromolecules. By Alexander McPherson. Cold Spring Harbor Laboratory Press. 1999. 586 pages. ISBN 0 87969 527 7. Price $97.00.

Alexander McPherson has been one of the leading lights in trying to systematize and de-mystify the process of obtaining protein crystals. This 500 page book gathers together the different approaches, theories, philosophies and folklore connected with what is still a suck-it-and see branch of science.

Genome sequencing operations are leaving a wash of new 'proteomic projects' in their wake. One such project involves networks of crystallographic groups tackling series of related proteins. A major long-term goal from this approach will be to obtain a library of protein structures which provide 3D templates for members of all the different structural families. Such projects will rapidly increase the current structural database from the current level of nearly 10000 entries to over 100000 within a few years. The strategy will be the same for each structure; identify and clone the gene; overexpress the protein; grow suitable protein crystals and solve the 3D X-ray structure. X-ray

crystallography is still the only technique that can routinely provide atomic resolution structures for large molecules of over 30000 Dalton (250 amino acids) and the prerequisite for such an analysis is to obtain a stable, single, well formed protein crystal with ideal dimensions of $0{\cdot}2 \times 0{\cdot}2 \times 0{\cdot}2$ mm. For smaller proteins, Nuclear Magnetic Resonance is still an option with the advantage that only soluble protein (and no crystals) is required, though both NMR and X-ray methods require milligram quantities of material.

The book on 'Crystallisation of Biological Macromolecules' is therefore very timely as the importance of the protein crystallization step is taking on a more critical role as more and more proteins are being studied. There is also a healthy trend of molecular biologists picking up X-ray crystallography as an additional skill. Or, put another way, the route to obtaining suitable crystals for X-ray analysis study is often achieved through molecular biology techniques. Crystallography and molecular biology overlap at the protein production and crystallisation stage and with this in mind, McPherson has attempted to cover most topics in a way that is accessible to those without expert knowledge in either area.

This lavishly produced book with lots of colour pictures has an almost coffee-table feel to it. It contains a host of interesting snippets of information and is well referenced. One of the most readable chapters for me was the 'History and Character of Macromolecular Crystals' chronicling the first published protein crystallization papers on haemoglobin (Hünfeld 1840) and the landmark studies by Sumner on concanavalin in 1919. The intriguing work of O. Füke (1850) involved crystallization of haemoglobin from man, horse, bullock, dog, cat, fish, pig and pigeon. This was followed by an amazing 500 page book published in 1909 which apparently described an attempt to use the crystal forms of haemoglobin from a variety of organisms to provide an alternative biological classification. This historical introductory chapter is followed by two chapters on 'Principles of Macromolecular Structure' and 'Purification and Characterisation of Biological Macromolecules' that make a brave attempt to cover these broad topics. They are informative, however, they probably require too much background information for the newcomer to the field and are not detailed enough for the experts looking for tricks, tips and new insights.

The real meat of the book is in the central six chapters discussing Principles, Procedures, Considerations, Strategies, Special Approaches and Mechanisms of Macromolecular Crystallisation. There is indeed a lot of useful information here covering all the major aspects of the subject. I felt, however, that the book would have benefited considerably from some serious editing. Reading the book from cover to cover in a few sessions induced a vague sensation of *déjà vu*; some topics are discussed or described in rather similar ways in two or three of the chapters. In one case, an identical figure (the phase diagram for lysozyme) is presented in chapter 4 and 6 and discussed in similar ways. Similarly, a bit of work on editing the figures would have helped readability. Some of the copious illustrations could have been helped by more complete legends, or discussed more thoroughly in the text.

The thermodynamics and kinetics of protein crystal growth are complex and rely on a host of parameters including protein concentration, precipitant concentration, pH and temperature. So it is perhaps reasonable to take a more qualitative and pictorial approach used in this book to describe this process. Current theory suggests that the first step in crystal growth is the formation of nuclei of a critical size in a supersaturated solution. The critical nuclei consist of a few hundred molecules for large proteins and a few thousand molecules for smaller proteins. The use of Atomic Force Microscopy (AFM) provides a clear picture of the subsequent development of crystal faces during crystal growth. A number of beautifully illustrated examples of crystal stacking and screw dislocation faults are provided in chapters 8 and 9. It is an interesting observation that in contrast to small molecule crystals, protein crystals grow more slowly and only reach a limited size. Most protein crystals have between 10^4 and 10^6 defects/cm^2 compared to a value of about 100/cm^2 for small molecule crystals and AFM evidence suggests that there may be a correlation between the size of the crystal, how well it diffracts X-rays and the defect density.

There is also a useful set of appendices at the end of the book which provides a compendium of screening conditions and some useful recipes for buffer preparation and other standard laboratory procedures. It is ironic that the new wave of 'structural proteomics' projects have a major bottleneck in the empirical and capricious science of crystal growth. This makes the compilation of approaches in this particularly timely providing a good and detailed coverage of the area. Despite my feeling that a leaner and more focused book could have been produced, I can strongly recommend it to novices and experts alike.

MALCOLM WALKINSHAW
Institute of Cell and Molecular Biology,
University of Edinburgh

https://doi.org/10.1017/S0016672399229072 Published online by Cambridge University Press